

PATENT ABSTRACTS OF JAPAN

(11) Publication number : 11-338836
 (43) Date of publication of application : 10.12.1999

(51) Int. Cl. G06F 15/16
 G06F 13/00

(21) Application number : 10-142655

(71) Applicant : NIPPON TELEGR & TELEPH CORP
 <NTT>

(22) Date of filing : 25.05.1998

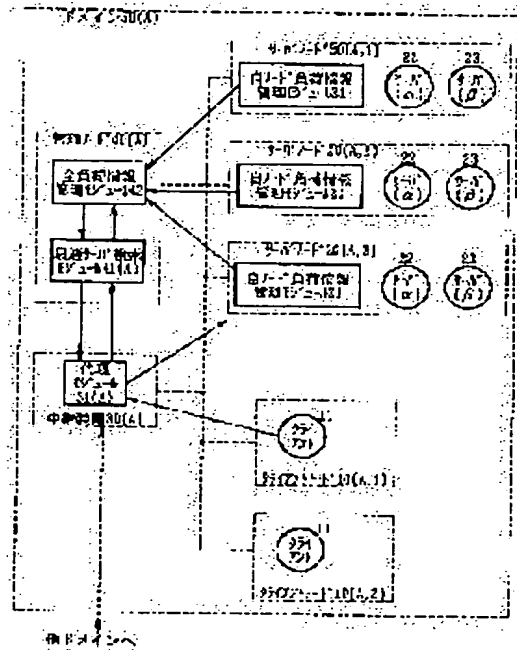
(72) Inventor : KATAYAMA MINORU

(54) LOAD DISTRIBUTION SYSTEM FOR COMPUTER NETWORK

(57) Abstract:

PROBLEM TO BE SOLVED: To improve responsiveness to a service request from a client to a server.

SOLUTION: This system is provided with a load monitoring means 21 by nodes for detecting the respective load levels of plural server nodes, optimum server retrieving means 40 for identifying the optimum server node based on the detected levels of loads out of plural server nodes, to which servers 22 and 23 capable of executing the service requested from a client node are connected, while grasping the kind of the service provided by the respective plural server nodes and the levels of the loads of respective server nodes detected by the load monitoring means 21 by nodes, and repeating control means 30 for repeating the service request from the client node to servers 22 and 23 and selectively sending the service request from the client node to the optimum server node.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C) ; 1998, 2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-338836

(43) 公開日 平成11年(1999)12月10日

(51) Int.Cl.⁶

G 0 6 F 15/16
13/00

識別記号

3 8 0
3 5 7

F I

G 0 6 F 15/16
13/00

3 8 0 Z
3 5 7 Z

審査請求 未請求 請求項の数 5 O L (全 13 頁)

(21) 出願番号

特願平10-142655

(22) 出願日

平成10年(1998) 5月25日

(71) 出願人 000004226

日本電信電話株式会社
東京都千代田区大手町二丁目3番1号

(72) 発明者 片山 稔

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(74) 代理人 弁理士 古谷 史旺

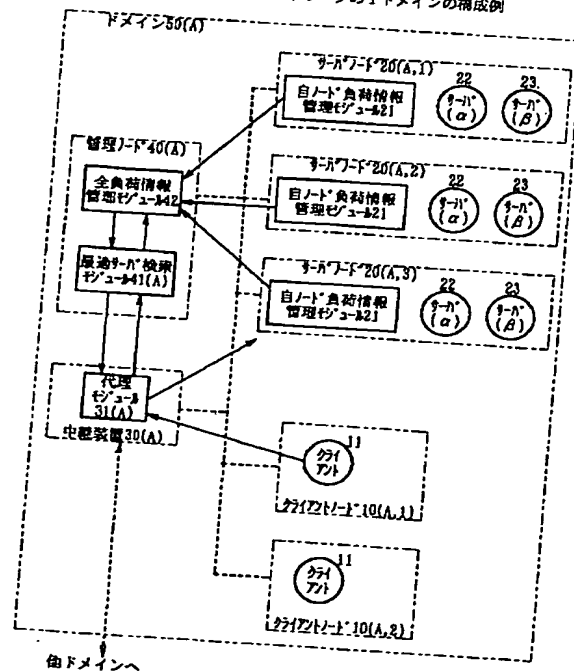
(54) 【発明の名称】 コンピュータネットワークの負荷分散システム

(57) 【要約】

【課題】 本発明はコンピュータネットワークの負荷分散システムにおいてクライアントからサーバへのサービス要求に対する応答性の改善を目的とする。

【解決手段】 複数のサーバノードの各々の負荷レベルを検出するノード別負荷監視手段21と複数のサーバノードの各々が提供するサービスの種類及びノード別負荷監視手段21が検出した各サーバノードの負荷のレベルを把握してクライアントノードから要求されるサービスを実行可能なサーバ22, 23が接続された複数のサーバノードの中から検出された負荷のレベルに基づいて最適サーバノードを識別する最適サーバ検索手段40とクライアントノードからサーバ22, 23へのサービス要求を中継するとともに最適サーバノードに対して選択的にクライアントノードからのサービス要求を送出する中継制御手段30とを設けた。

実施の形態の負荷分散システムを創る
コンピュータネットワークの1ドメインの構成例



【特許請求の範囲】

1

【請求項 1】 複数のサーバノードとクライアントノードとが形成され、前記複数のサーバノードのそれぞれには少なくとも 1 つのサーバが接続されたコンピュータネットワークの負荷分散システムにおいて、前記複数のサーバノードのそれぞれについて負荷のレベルを検出するノード別負荷監視手段と、

前記複数のサーバノードのそれぞれが提供するサービスの種類及び前記ノード別負荷監視手段が検出した各サーバノードの負荷のレベルを把握して、前記クライアントノードから要求されるサービスを実行可能なサーバが接続された複数のサーバノードの中から、検出された負荷のレベルに基づいて最適サーバノードを識別する最適サーバ検索手段と、

前記クライアントノードからサーバへのサービス要求を中継するとともに、前記最適サーバ検索手段が識別した最適サーバノードに対して選択的に前記クライアントノードからのサービス要求を送出する中継制御手段とを設けたことを特徴とするコンピュータネットワークの負荷分散システム。

【請求項 2】 請求項 1 記載のコンピュータネットワークの負荷分散システムにおいて、所定時間内に所定以上の負荷の増大が検出されたサーバノードに対しては、前記クライアントノードからサーバへのサービスの割り当てを一時的に禁止する割り当て制御手段を設けたことを特徴とするコンピュータネットワークの負荷分散システム。

【請求項 3】 請求項 2 記載のコンピュータネットワークの負荷分散システムにおいて、前記割り当て制御手段が、所定時間内に所定以上の負荷の増大が検出されたサーバノードについては、前記ノード別負荷監視手段が検出したサーバノードの負荷のレベルを予め定めた最大値に強制的に変更することを特徴とするコンピュータネットワークの負荷分散システム。

【請求項 4】 複数のドメインで構成され、各々のドメインには複数のサーバノードとクライアントノードとが形成され、前記複数のサーバノードのそれぞれには少なくとも 1 つのサーバが接続されたコンピュータネットワークの負荷分散システムにおいて、

前記複数のサーバノードのそれぞれについて負荷のレベルを検出するノード別負荷監視手段と、

複数のドメインの間の通信にかかる第 1 の通信所要時間及び各ドメイン内の所定位置と各サーバノードとの間の通信にかかる第 2 の通信所要時間を検出する通信時間検出手段と、

複数のドメインの前記複数のサーバノードのそれぞれが提供するサービスの種類、前記ノード別負荷監視手段が検出した各サーバノードの負荷のレベル、前記通信時間検出手段が検出した第 1 の通信所要時間及び第 2 の通信所要時間に基づいて、前記クライアントノードから要求さ

10 【請求項 5】 請求項 4 記載のコンピュータネットワークの負荷分散システムにおいて、各ドメイン内の前記中継制御手段と各サーバノードとの間の通信経路の距離をほぼ同一に定め、前記通信時間検出手段が、前記中継制御手段と各サーバノードとの間で実際に伝送された信号の往復所要時間に基づいて前記第 2 の通信所要時間を検出することを特徴とするコンピュータネットワークの負荷分散システム。

【発明の詳細な説明】

【0001】

20 【発明の属する技術分野】本発明は、複数のサーバノードとクライアントノードとが形成され、前記複数のサーバノードのそれぞれには少なくとも 1 つのサーバが接続されたコンピュータネットワークの負荷分散システムに関する。

【0002】

30 【従来の技術】例えば、ユーザが要求するアプリケーションプログラムを実行可能なサーバが複数接続されたコンピュータネットワークにおいては、ユーザの要求に応じてアプリケーションプログラムを実行するサーバを適宜選別することにより、サーバの負荷を分散させることができる。

40 【0003】特定のサーバに処理が集中し前記サーバの負荷が過大になると、処理に遅延が生じるのでサービスのリアルタイム性が失われる。ユーザの要求に応じて実行される様々なサービスを複数のサーバに適宜に分散させることにより、処理の集中を防止し、サービスの遅延を避けることができる。従来のコンピュータネットワークの負荷分散システムは、図 7 に示すように構成されている。このコンピュータネットワークは、複数のクライアントノードと複数のサーバノードとで構成されている。各々のサーバノードには、その負荷レベルを計測して管理する自ノード負荷情報管理モジュールが設けられ、各々のクライアントノードには、各サーバノードの負荷レベルを収集して管理する他ノード負荷情報管理モジュールが設けられる。

50 【0004】各々のクライアントノードの他ノード負荷情報管理モジュールは、全てのサーバノードで検出された負荷レベルのデータをそれぞれ受け取って、負荷の小さいサーバノードを検出するために利用する。

【0005】

【発明が解決しようとする課題】従来のコンピュータネットワークの負荷分散システムにおいては、それぞれのサーバノードは、全てのクライアントノードに対して自ノードの負荷レベルのデータを通知する必要がある。従って、クライアントノード数の増大に伴って、サーバノードの負荷とネットワークの負荷が重くなってしまう。

【0006】サーバノードとクライアントノードとの間では、所定の時間間隔で定期的に負荷レベルのデータが転送される。この時間間隔が比較的大きい場合には、制御の遅れが生じる可能性がある。例えば、サーバノードの負荷レベルが急速に上昇すると、負荷の増大をクライアントノードが認識する前にサーバノードの負荷が過大になり、負荷が過大なサーバノードに対して更にサービスを割り当てるため処理の遅延が発生する。

【0007】ところで、ローカルエリアネットワーク(LAN)などのドメインを複数接続したコンピュータネットワークにおいては、互いに異なるドメインに属するサーバノードも利用できる。しかし、一般にドメイン間では通信経路の距離が長いので、通信所要時間が長くなる。負荷の大きさだけを考慮して他のドメインに属するサーバノードを利用すると、サーバノードからの応答時間が遅くなり、応答までの時間の予測もできない場合が多い。

【0008】人とノードとの間でインタラクティブな処理が行われるアプリケーションに対しては、サービスを提供するサーバに対して数秒から数十秒程度の応答性が必要とされる。本発明は、コンピュータネットワークの負荷分散システムにおいて、クライアントからサーバへのサービス要求に対する応答性を改善することを目的とする。

【0009】

【課題を解決するための手段】請求項1のコンピュータネットワークの負荷分散システムは、複数のサーバノードとクライアントノードとが形成され、前記複数のサーバノードのそれぞれには少なくとも1つのサーバが接続されたコンピュータネットワークの負荷分散システムにおいて、前記複数のサーバノードのそれぞれについて負荷のレベルを検出するノード別負荷監視手段と、前記複数のサーバノードのそれぞれが提供するサービスの種類及び前記ノード別負荷監視手段が検出した各サーバノードの負荷のレベルを把握して、前記クライアントノードから要求されるサービスを実行可能なサーバが接続された複数のサーバノードの中から、検出された負荷のレベルに基づいて最適サーバノードを識別する最適サーバ検索手段と、前記クライアントノードからサーバへのサービス要求を中継するとともに、前記最適サーバ検索手段が識別した最適サーバノードに対して選択的に前記クライアントノードからのサービス要求を送出する中継制御手段とを設けたことを特徴とする。

【0010】ノード別負荷監視手段は、複数のサーバノ

ードのそれぞれについて負荷のレベルを検出する。最適サーバ検索手段は、複数のサーバノードのそれぞれが提供するサービスの種類及び前記ノード別負荷監視手段が検出した各サーバノードの負荷のレベルを把握して、前記クライアントノードから要求されるサービスを実行可能なサーバが接続された複数のサーバノードの中から、検出された負荷のレベルに基づいて、例えば負荷のレベルが最小の最適サーバノードを識別する。

【0011】中継制御手段は、クライアントノードからサーバへのサービス要求を中継するとともに、前記最適サーバ検索手段が識別した最適サーバノードに対して選択的に前記クライアントノードからのサービス要求を送出する。本発明においては、各サーバノードの負荷のレベルが最適サーバ検索手段によって集中的に管理されるので、各サーバノードはクライアントノードに対して負荷レベルのデータを通知する必要がある。従って、クライアントノード数が増大しても、負荷レベルのデータの通知のために各サーバノードの処理が増大することがなく、ネットワーク上のトラヒックも増大しない。

【0012】また、検出された負荷のレベルが最小のサーバノードに対して選択的に前記クライアントノードからのサービス要求を送出するので、このサービス要求に対して高い応答性が得られる。請求項2は、請求項1記載のコンピュータネットワークの負荷分散システムにおいて、所定時間内に所定以上の負荷の増大が検出されたサーバノードに対しては、前記クライアントノードからサーバへのサービスの割り当てを一時的に禁止する割り当て制御手段を設けたことを特徴とする。

【0013】ノード別負荷監視手段から最適サーバ検索手段に対して負荷のレベルを通知する時間間隔が比較的大きい場合には、サーバノードの負荷レベルが急速に上昇すると、負荷の増大を最適サーバ検索手段が認識する前にサーバノードの負荷が過大になり、負荷が過大なサーバノードに対して更にサービスを割り当てる可能性がある。また、負荷のレベルを通知する時間間隔を小さくすると、その処理に要する負担やネットワークのトラヒックが増大する。

【0014】割り当て制御手段は、所定時間内に所定以上の負荷の増大が検出されたサーバノードに対しては、前記クライアントノードからサーバへのサービスの割り当てを一時的に禁止する。従って、負荷レベルが上昇傾向にあり、負荷レベルが過大になる可能性の高いサーバノードに対しては、負荷レベルが過大になる前にサービスの割り当てが禁止される。

【0015】このため、ノード別負荷監視手段から最適サーバ検索手段に対して負荷のレベルを通知する時間間隔が比較的大きい場合であっても、負荷が過大なサーバノードに対してサービスを割り当てることを未然に防止できる。請求項3は、請求項2記載のコンピュータネットワークの負荷分散システムにおいて、前記割り当て制

御手段が、所定時間内に所定以上の負荷の増大が検出されたサーバノードについては、前記ノード別負荷監視手段が検出したサーバノードの負荷のレベルを予め定めた最大値に強制的に変更することの特徴とする。

【0016】所定時間内に所定以上の負荷の増大が検出されたサーバノードについては、前記ノード別負荷監視手段から最適サーバ検索手段に通知される負荷のレベルが最大値、すなわち過大レベルになるため、負荷が急速に上昇するサーバノードが最適サーバとして認識されることがなくなる。請求項4のコンピュータネットワークの負荷分散システムは、複数のドメインで構成され、各々のドメインには複数のサーバノードとクライアントノードとが形成され、前記複数のサーバノードのそれぞれには少なくとも1つのサーバが接続されたコンピュータネットワークの負荷分散システムにおいて、前記複数のサーバノードのそれぞれについて負荷のレベルを検出するノード別負荷監視手段と、複数ドメインの間の通信にかかる第1の通信所要時間及び各ドメイン内の所定位置と各サーバノードとの間の通信にかかる第2の通信所要時間を検出する通信時間検出手段と、複数ドメインの前記複数のサーバノードのそれぞれが提供するサービスの種類、前記ノード別負荷監視手段が検出した各サーバノードの負荷のレベル、前記通信時間検出手段が検出した第1の通信所要時間及び第2の通信所要時間に基づいて、前記クライアントノードから要求されるサービスを実行可能なサーバが接続されたサーバノードの中から1つのサーバノードを最適サーバノードとして識別する最適サーバ検索手段と、前記クライアントノードからサーバへのサービス要求を中継するとともに、前記最適サーバ検索手段が識別した最適サーバノードに対して選択的に前記クライアントノードからのサービス要求を送出する中継制御手段とを設けたことを特徴とする。

【0017】ノード別負荷監視手段は、複数のサーバノードのそれぞれについて負荷のレベルを検出する。通信時間検出手段は、複数ドメインの間の通信にかかる第1の通信所要時間及び各ドメイン内の所定位置と各サーバノードとの間の通信にかかる第2の通信所要時間を検出する。最適サーバ検索手段は、複数ドメインの前記複数のサーバノードのそれぞれが提供するサービスの種類、前記ノード別負荷監視手段が検出した各サーバノードの負荷のレベル、前記通信時間検出手段が検出した第1の通信所要時間及び第2の通信所要時間に基づいて、前記クライアントノードから要求されるサービスを実行可能なサーバが接続されたサーバノードの中から1つのサーバノードを最適サーバノードとして識別する。

【0018】中継制御手段は、前記クライアントノードからサーバへのサービス要求を中継するとともに、前記最適サーバ検索手段が識別した最適サーバノードに対して選択的に前記クライアントノードからのサービス要求を送出する。最適サーバ検索手段は、ネットワークに接

続された複数ドメインの中から最適サーバノードを識別するので、サービスを要求するクライアントノードと同一のドメイン内に負荷の小さいサーバノードが存在しない場合には、他のドメインに属する負荷の小さいサーバノードに対してサービスを割り当てることができる。

【0019】また、最適サーバ検索手段は、各サーバノードの負荷のレベルだけでなく、前記第1の通信所要時間及び第2の通信所要時間を考慮して最適サーバノードを識別するので、クライアントノードから距離の離れたドメインに属するサーバノードについては、負荷のレベルが小さくても頻繁にサービスを割り当てることはいない。

【0020】つまり、負荷のレベルが同一であれば、クライアントノードに最も近いドメインに属するサーバノードに優先的にサービスを割り当てることができるので、効率的に負荷を分散できる。請求項5は、請求項4記載のコンピュータネットワークの負荷分散システムにおいて、各ドメイン内の前記中継制御手段と各サーバノードとの間の通信経路の距離をほぼ同一に定め、前記通信時間検出手段が、前記中継制御手段と各サーバノードとの間で実際に伝送された信号の往復所要時間に基づいて前記第2の通信所要時間を検出することを特徴とする。

【0021】通信時間検出手段は、前記中継制御手段と各サーバノードとの間で実際に伝送された信号の往復所要時間に基づいて前記第2の通信所要時間を検出する。各ドメイン内の前記中継制御手段と各サーバノードとの間の通信経路の距離がほぼ同一であるため、基本的に同一ドメイン内では各サーバノードの通信時間には差が生じない。従って、同一ドメイン内の第2の通信所要時間に含まれる誤差は、最適サーバノードの識別に悪影響を及ぼさない。

【0022】

【発明の実施の形態】実施の形態の負荷分散システムの構成及び動作を図1～図6に示す。この形態は全ての請求項に対応する。図1はこの形態の負荷分散システムを備えるコンピュータネットワークの1ドメインの構成例を示すブロック図である。図2はこの形態の負荷分散システムを備えるコンピュータネットワークの2ドメインの構成例を示すブロック図である。

【0023】図3は自ノード負荷情報管理モジュールの負荷通知処理の内容を示すフローチャートである。図4は最適サーバ検索モジュールの検索処理の内容を示すフローチャートである。図5はドメイン内の距離測定処理の内容を示すフローチャートである。図6はドメイン間の距離測定処理の内容を示すフローチャートである。この形態では、請求項1のノード別負荷監視手段、最適サーバ検索手段及び中継制御手段は、それぞれ自ノード負荷情報管理モジュール21、管理ノード40及び中継装置30に対応する。

【0024】請求項2及び請求項3の割り当て制御手段は、自ノード負荷情報管理モジュール21に対応する。請求項4のノード別負荷監視手段、通信時間検出手段、最適サーバ検索手段及び中継制御手段は、それぞれ自ノード負荷情報管理モジュール21、自ノード負荷情報管理モジュール21、管理ノード40及び中継装置30に対応する。

【0025】本発明の負荷分散システムは、図1及び図2に示すようなコンピュータネットワークで実施される。この種のコンピュータネットワークは、図1に示すような1つのドメイン50(A)だけで構成される場合と、図2に示すように互いに通信可能な2つ以上のドメイン50(A)、50(B)、・・・で構成される場合とがある。図1及び図2においては、各構成要素の符号に付加した括弧内の記号A、Bでそれが属するドメインの区分を表し、同様に括弧内の番号1、2、3によってノードの区分を表す。以下の説明においても、必要に応じてこれらの区分を各構成要素の符号に付けて示す。

【0026】図1を参照すると、このドメイン50には、複数のクライアントノード10、複数のサーバノード20、中継装置30及び管理ノード40が備わっている。このドメイン50は、現実的にはローカルエリアネットワーク(LAN)のように、多数のノードをハブやルータのようなスイッチを介してネットワークで互いに接続される。

【0027】図1においては、複数のクライアントノード10、複数のサーバノード20、中継装置30及び管理ノード40が、点線で示すようにネットワークで互いに接続されている。図1及び図2に示す実線の矢印は、ノード割当制御における主な信号の流れを示している。なお、実際には多数のクライアントノード10及び多数のサーバノード20がネットワークに接続されるが、図1及び図2にはそれらの一部分だけが示されている。

【0028】中継装置30は、交換機と同様の構成要素を備えており、ドメイン50内の各構成要素及び複数ドメイン50間を必要に応じて接続し、接続した構成要素の間の通信を可能にする。中継装置30には代理モジュール31が備わっている。この代理モジュール31は、例えば中継装置30に内蔵されたコンピュータが実行可能なプログラムであり、後述する所定の機能を実現する。

【0029】サーバノード20には、1つ以上のサーバ22、23が必要に応じて接続される。サーバ22、23は、サーバノード20に内蔵された1つ以上のコンピュータと、それによって実行されるアプリケーションのサーバプログラムとで構成される。この例では、記号α、βで区分される互いに種類の異なるアプリケーションのサーバプログラムがサーバ22及び23に備わっている。

【0030】クライアントノード10には、1つ以上の

クライアント11が必要に応じて接続される。クライアント11は、クライアントノード10に内蔵された1つ以上のコンピュータと、それによって実行されるアプリケーションのクライアントプログラムとで構成される。クライアント11がある種のアプリケーションを利用する場合には、クライアント11におけるクライアントプログラムからの要求に応じて、何れかのサーバノード20に属するサーバ22又は23のサーバプログラムが起動する。サーバ22又は23のサーバプログラムは、所定のサービスを実行してその結果をクライアント11に返す。

【0031】図1に示すドメイン50においては、記号αに対応する種類のアプリケーションを実行するサーバ22が複数のサーバノード20に接続されているので、クライアント11からの要求に応じたサービスを複数のサーバノード20(A,1)、20(A,2)、20(A,3)の何れでも処理できる。この形態では、クライアント11が直接何れかのサーバノード20を選択することはない。この形態では、負荷を分散するためにクライアント11からのサービス要求は中継装置30に入力され、中継装置30が複数のサーバノード20(A,1)、20(A,2)、20(A,3)の何れかを選択して、選択された1つのサーバノード20に属するサーバ22に対してサービス要求を中継する。

【0032】各サーバノード20には、そのノード内部における負荷の大きさを管理するために自ノード負荷情報管理モジュール21が設けてある。この自ノード負荷情報管理モジュール21は、サーバノード20に内蔵されたコンピュータで実行される制御プログラムである。管理ノード40は、中継装置30からの要求に応じて、ドメイン50内部における最適なサーバノード20を決定し、その結果を中継装置30に通知する。管理ノード40には、最適サーバ検索モジュール41及び全負荷情報管理モジュール42が備わっている。最適サーバ検索モジュール41及び全負荷情報管理モジュール42は、管理ノード40に内蔵されたコンピュータで実行される制御プログラムである。

【0033】全負荷情報管理モジュール42は、ドメイン50内の全てのサーバノード20に関する負荷情報を各サーバノード20の自ノード負荷情報管理モジュール21から受け取って管理する。最適サーバ検索モジュール41は、ドメイン50(A)内の全負荷情報管理モジュール42が管理する各サーバノード20(A)の負荷情報と他のドメイン50(B)が管理する各サーバノード20(B)の負荷情報及び距離情報に基づいて、負荷の分散と処理のリアルタイム性の観点から最適サーバノードを決定する。

【0034】ドメイン50内では、中継装置30と各々のサーバノード20との間の距離が互いにほぼ同一になるように各構成要素が配置されている。このため、中継

装置30と各サーバノード20との間の通信所要時間はほぼ同じになる。従って、ドメイン50内における最適サーバノードは、負荷が最小のサーバノード20に決定される。勿論、要求されたサービスを実行可能なサーバ22又は23が備わったサーバノード20に限定される。

【0035】例えば、何れかのクライアント11が中継装置30の代理モジュール31にサーバ(α)22で実行可能な処理の要求を送出すると、代理モジュール31は最適サーバ検索モジュール41に、負荷とリアルタイム性の観点において最適なサーバ(α)22の検索を依頼する。最適サーバ検索モジュール41は、最適なサーバ(α)22の検索依頼を受けると、全負荷情報管理モジュール42に対して、サーバ(α)22が属するサーバノード20の負荷情報を要求する。

【0036】全負荷情報管理モジュール42は、常に各サーバノード20の自ノード負荷情報管理モジュール21からの負荷情報を管理しているので、最適サーバ検索モジュール41からサーバノードの負荷情報を要求されると、管理している負荷情報を最適サーバ検索モジュール41に返却する。最適サーバ検索モジュール41は、全負荷情報管理モジュール42から負荷情報を受け取って、負荷とリアルタイム性の観点において最適なサーバ(α)22を検出し、それが属するサーバノード20及びサーバ(α)22の識別子とその位置を示す情報を返却する。

【0037】各サーバノード20に備わった自ノード負荷情報管理モジュール21の負荷通知処理の各ステップの内容について図3を参照して説明する。ステップS11では、このプログラムを実行するコンピュータの内部メモリに割り当てたカウンタ*i*及びレジスタ*X*の値をクリアする。

【0038】ステップS12では、自ノードの負荷の大きさを求める。具体的には、自ノードに接続されたコンピュータについて検出された最新のCPU利用率、メモリの利用率及びネットワークの負荷状態(そのノードに接続されているクライアント数)をパラメータとして予め定めた計算式から負荷の大きさを計算する。この例では、負荷の大きさは1から予め定めた最大値*X_m*までの*X_m*段階のレベルに分類して表される。なお、負荷のレベルの範囲はノード毎に独立に定めても良い。例えば、負荷のレベルの最小値をCPUの能力、メモリ量及びネットワークの帯域に応じて、 X_1 ($1 < X_1 < X_m$)に変更しても良い。

【0039】ステップS12は一定の時間周期で繰り返し実行される。ステップS12で求められた負荷の大きさは、カウンタ*i*の値によって特定されるレジスタ*X*(*i*)に保持される。カウンタ*i*の値は、ステップS12を実行する度にステップS18又はS11で更新される。ステップS13では、ステップS12の結果が保持

されるレジスタ*X*(*i*)の内容を最大値*X_m*と比較する。ステップS12で検出された負荷の大きさが最大値*X_m*未満であればステップS14に進み、検出された負荷の大きさが最大値*X_m*であればステップS19に進む。

【0040】ステップS14では、(*i*-1)で特定されるレジスタ*X*(*i*-1)が保持する値と予め定めた閾値*X_a*とを比較する。つまり、1周期前にステップS12で検出された負荷の大きさを閾値*X_a*と比較する。レジスタ*X*(*i*-1)の値が閾値*X_a*未満であればステップS16に進み、そうでなければステップS15に進む。ステップS15では、レジスタ*X*(*i*)が保持する値とレジスタ*X*(*i*-1)が保持する値との差分を予め定めた閾値*X_b*と比較する。つまり、負荷の1検出周期の間の負荷の変化量を閾値*X_b*と比較する。前記変化量が閾値*X_b*未満であればステップS16に進み、それ以外の場合にはステップS20に進む。例えば、負荷が急速に増大する場合にはステップS15からS20に進む可能性がある。

【0041】ステップS16では、レジスタ*X*(*i*)が保持する値をレジスタ*X*に加算してレジスタ*X*の値を更新する。ステップS17では、カウンタ*i*の値を(*k*-1)と比較する。*k*は定数である。カウンタ*i*の値が(*k*-1)未満ならステップS18に進み、それ以外の場合にはステップS21に進む。

【0042】ステップS18では、カウンタ*i*の値を更新する。ステップS19では、レジスタ*N_p*が保持する値を最大値*X_m*と比較する。レジスタ*N_p*が保持する値は、以前にステップS22の処理で全負荷情報管理モジュール42に通知した負荷情報の値である。レジスタ*N_p*の値が最大値*X_m*未満の場合にはステップS20に進み、それ以外の場合にはステップS11に進む。

【0043】ステップS20では、レジスタ*N*に最大値*X_m*を保持する。ステップS21では、レジスタ*X*の値を定数*k*で除算した結果をレジスタ*N*に保持する。ステップS22では、レジスタ*N*が保持する値を検出された負荷のレベルとして全負荷情報管理モジュール42に通知する。

【0044】ステップS23では、通知した負荷のレベルをレジスタ*N_p*に保存する。従って、ステップS12で検出した負荷のレベルが比較的小さい場合には、ステップS12で検出された負荷のレベルがレジスタ*X*に累積される。そして、負荷のレベルを*k*回検出する度に、検出された*k*個のデータの累積値の平均値がステップS21で求められる。この平均値は、所定周期(例えば1分間)の負荷のレベルの平均値である。この平均値が前記所定周期毎にステップS22で全負荷情報管理モジュール42に通知される。

【0045】但し、ステップS12で検出した負荷のレベルが最大値*X_m*と等しく、前回のステップS22で全負荷情報管理モジュール42に通知した負荷のレベルが

最大値Xmより小さい場合には、ステップS13-S19-S20-S22を実行するので直ちに最大値Xmが全負荷情報管理モジュール42に通知される。

【0046】また、ステップS12で1周期前に検出された負荷のレベルが閾値Xa以上で、しかも負荷検出周期の1周期の間の負荷のレベルの変化量が閾値Xb以上になると、ステップS14-S15-S20-S22を実行するので直ちに最大値Xmが全負荷情報管理モジュール42に通知される。つまり、急速な負荷の増大を検出した場合には、そのサーバノード20に対する更なる処理の割当を一時的に禁止するために、負荷のレベルとして実際に検出された値より大きな最大値Xmを通知する。

【0047】各サーバノード20の自ノード負荷情報管理モジュール21が図3に示すような処理を実行するので、管理ノード40の全負荷情報管理モジュール42には、ドメイン50内の全てのサーバノード20から、その負荷のレベルを示す情報が集まる。次に、最適サーバ検索モジュール41によって処理される検索処理の内容について図4を参照して説明する。また、ドメイン50(A)に属するクライアント11がサービスの割当を要求する場合の例について具体的に説明する。

【0048】ステップS31では、検索要求があるまで待機する。クライアント11がサービス割当を中継装置30の代理モジュール31に要求した場合には、代理モジュール31は同じドメイン50に属する最適サーバ検索モジュール41に対して、検索要求を送出する。また、図2に示すように複数ドメイン50(A)、50(B)が接続されている場合にドメイン50(A)に属する最適サーバ検索モジュール41(A)が検索要求を受けると、最適サーバ検索モジュール41(A)は、中継装置30(A)の代理モジュール31(A)及び中継装置30(B)の代理モジュール31(B)を介して、ドメイン50(B)に属する最適サーバ検索モジュール41(B)に対して検索要求を送出する。

【0049】検索要求を受けると、最適サーバ検索モジュール41の処理はステップS31からS32に進む。ステップS32では、クライアント11から要求されたサービスの種類を示す識別子の情報に基づいて、そのサービスを実行可能なサーバ22(又は23)が接続された全てのサーバノード20を識別する。

【0050】例えば、図2においてクライアント11が記号αで表されるサービスの割当を要求した場合には、最適サーバ検索モジュール41はサーバ(α)22を備えるサーバノード20(A,1)、20(A,2)、20(A,3)の各々の識別子をステップS32で取得する。ステップS33では、他のドメインが利用可能か否かを識別する。図2に示すように複数のドメイン50(A)、50(B)が互いに接続されている場合には、ステップS33からS34に進む。

【0051】ステップS34では、他のドメイン50に属する最適サーバ検索モジュール41に対して検索すべきサーバ22、23の種類(α、β)を示す識別子とともに検索要求を送出する。最適サーバ検索モジュール41(A)から検索要求を受けた最適サーバ検索モジュール41(B)も、最適サーバ検索モジュール41(A)と同様に図4の処理を実行する。

【0052】ステップS35では、ステップS32で検出されたサーバノード20が1つか複数かを識別する。複数のサーバノード20が検出された場合には、ドメイン50内の最適なサーバノード20を選択するためにステップS35からS36に進む。ステップS36では、クライアント11からの要求を実行可能なサーバ22

(又は23)が接続された全てのサーバノード20の負荷情報(負荷のレベル)を全負荷情報管理モジュール42から取得する。

【0053】ステップS37では、クライアント11からの要求を実行可能なサーバ22(又は23)が接続された全てのサーバノード20の中で負荷のレベルが最小の1つのサーバノード20を決定する。ここで決定されるサーバノード20は、ドメイン50内における最適なサーバノード20である。なお、負荷のレベルが最小のサーバノード20が複数存在する場合には、その中の1つをランダムに選択する。

【0054】最適サーバ検索モジュール41(A)がステップS34で他ドメイン50(B)の最適サーバ検索モジュール41(B)に検索を要求した場合には、最適サーバ検索モジュール41(B)の検索の結果をステップS38で取得する。最適サーバ検索モジュール41(B)の検索の結果は、中継装置30(B)の代理モジュール31(B)、中継装置30(A)の代理モジュール31(A)を介して最適サーバ検索モジュール41(A)に入力される。

【0055】最適サーバ検索モジュール41は、ドメイン50内の各サーバノード20と中継装置30との間の通信所要時間T1、並びに互いに隣接するドメイン50(A)、50(B)に属する隣接する中継装置30(A)、30(B)の間の通信所要時間T2を距離情報として管理している。ステップS38においては、他ドメイン50(B)に属する最適サーバ検索モジュール41(B)から、ドメイン50(B)に属するサーバノード20(B)の中で負荷のレベルが最小の1つのサーバノード20(B)を示す識別子と、該サーバノード20(B)と中継装置30(B)との間の通信所要時間T1(B)のデータとが、最適サーバ検索モジュール41(A)に到来する。

【0056】ステップS39では、最適サーバ検索モジュール41(A)は、ステップS37で検出されたドメイン50(A)内の最適な1つのサーバノード20(A)と、他ドメイン50(B)で検出された最適な1つのサーバノード20(B)等について、それらの負荷のレベルと距離の情報とを考慮して、ネットワーク全体として最適な1つ

のサーバノード20を決定する。

【0057】ドメイン50(A)内の最適な1つのサーバノード20(A)については、その距離として最適サーバ検索モジュール41が管理するサーバノード20(A)と中継装置30(A)との間の通信所要時間T1(A)を適用する。また、他ドメイン50(B)のサーバノード20(B)の距離としては、最適サーバ検索モジュール41(B)から受け取ったサーバノード20(B)と中継装置30(B)との間の通信所要時間T1(B)に、最適サーバ検索モジュール41(A)が管理するドメイン50をまたがる中継装置30(A)、30(B)の間の通信所要時間T2を加算した結果を適用する。

【0058】なお、複数のドメイン50をまたがるサーバノード20を利用する場合には、それぞれのドメイン50間の通信所要時間を累積することにより、通信所要時間T2が得られる。ステップS39では、所定の計算式を用いて、複数ドメイン50で検出された複数のサーバノード20の中から、最適な1つのサーバノード20を識別する。例えば、ドメイン50(A)内のサーバノード20(A)に比べて他ドメイン50(B)のサーバノード20(B)の負荷のレベルが小さい場合であっても、中継装置30(A)、30(B)の間の通信所要時間T2が大きい場合には、ドメイン50(A)内のサーバノード20(A)が最適サーバノードとして識別される。

【0059】ステップS40では、例えば検索を要求した依頼元の代理モジュール31に対して検索の結果を返却する。最適サーバ検索モジュール41(A)から検索を要求された他ドメイン50(B)に属する最適サーバ検索モジュール41(B)の場合には、ステップS40で依頼元の最適サーバ検索モジュール41(A)に中継装置30(B)及び30(A)を介して結果を通知する。

【0060】検索要求に対して最適サーバ検索モジュール41から検索結果を受け取った中継装置30の代理モジュール31は、最適サーバ検索モジュール41が決定した最適サーバノードに対して、クライアント11からのサービスを割り当てるように要求を中継する。

【0061】ドメイン50内におけるサーバノード20と中継装置30との間の通信所要時間T1は、図5に示す処理によって取得される。図5に示す処理は、例えばサーバノード20の電源が投入されたときに実行される。図5の処理について以下に説明する。サーバノード20は、ステップS51でサーバノード20に内蔵されたコンピュータの時計回路の時刻を参照して現在時刻t11を検出する。その直後に、ステップS52でデータ長の小さい所定のデータをサーバノード20からドメイン50内の中継装置30に送出する。

【0062】中継装置30の処理は、ドメイン50内のサーバノード20から送出された前記データを受信すると、ステップS61からS62に進む。ステップS62では、受信したデータを直ちに送信元のサーバノード2

0に折り返し送信する。サーバノード20は、中継装置30からの返信データを受信すると、ステップS53からS54に進む。ステップS54では、再びコンピュータに内蔵された時計回路の時刻を参照して現在時刻t12を検出する。

【0063】サーバノード20は、ステップS55で次の計算式に基づき通信所要時間T1を算出する。

$$T1 = (t12 - t11) / 2$$

サーバノード20は、求められた通信所要時間T1を次のステップS56で管理ノード40に対して送出する。この通信所要時間T1は、管理ノード40の最適サーバ検索モジュール41で管理される。

【0064】一方、ドメイン50をまたがる中継装置30(A)、30(B)の間の通信所要時間T2は、図6に示す処理によって取得される。図6に示す処理は、例えば中継装置30の電源が投入されたときに実行される。図6の処理について以下に説明する。中継装置30(A)は、ステップS71でそれに内蔵されたコンピュータの時計回路の時刻を参照して現在時刻t21を検出する。その直後に、ステップS72でデータ長の小さい所定のデータを中継装置30(A)から他ドメイン50の中継装置30(B)に送出する。

【0065】中継装置30(B)の処理は、中継装置30(A)から送出された前記データを受信すると、ステップS81からS82に進む。ステップS82では、受信したデータを直ちに送信元の中継装置30(A)に折り返し送信する。中継装置30(A)は、中継装置30(B)からの返信データを受信すると、ステップS73からS74に進む。ステップS74では、再びコンピュータに内蔵された時計回路の時刻を参照して現在時刻t22を検出する。

【0066】中継装置30(A)は、ステップS75で次の計算式に基づき通信所要時間T2を算出する。

$$T2 = (t22 - t21) / 2$$

中継装置30(A)は、求められた通信所要時間T2を次のステップS76でドメイン50内の管理ノード40(A)に対して送出する。この通信所要時間T2は、管理ノード40(A)の最適サーバ検索モジュール41(A)で管理される。

【0067】ドメイン50(B)においても、同様の処理を実行するので最適サーバ検索モジュール41(B)は通信所要時間T2を取得して管理できる。

【0068】

【発明の効果】本発明によれば、各サーバノードの負荷のレベルが最適サーバ検索手段によって集中的に管理されるので、各サーバノードはクライアントノードに対して負荷レベルのデータを通知する必要がない。従って、クライアントノード数が増大しても、負荷レベルのデータの通知のために各サーバノードの処理が増大することがなく、ネットワーク上のトラヒックも増大しない。

【0069】また、検出された負荷のレベルが最小のサーバノードに対して選択的に前記クライアントノードからのサービス要求を送出するので、このサービス要求に対して高い応答性が得られる。また、負荷が急速に増大するサーバノードに対してサービスの割り当てを一時的に禁止する割り当て制御手段を設けることにより、負荷レベルが過大になる可能性の高いサーバノードに対しては、負荷レベルが過大になる前にサービスの割り当てが禁止される。このため、ノード別負荷監視手段から最適サーバ検索手段に対して負荷のレベルを通知する時間間隔が比較的大きい場合であっても、負荷が過大なサーバノードにサービスを割り当てることを未然に防止できる。

【0070】また、複数ドメインを利用できる環境においては、ノード別負荷監視手段が検出した各サーバノードの負荷のレベルだけでなく、通信時間検出手段が検出した第1の通信所要時間及び第2の通信所要時間を考慮して最適サーバノードを識別する最適サーバ検索手段を設けることにより、サーバからの応答性を考慮して適度に負荷を分散できる。

【0071】また、各ドメイン内の前記中継制御手段と各サーバノードとの間の通信経路の距離をほぼ同一に定めることにより、検出された第2の通信所要時間に含まれる多少の誤差は許容できるので、第2の通信所要時間の検出が容易になる。更に、中継制御手段と各サーバノードとの間で実際に伝送された信号の往復所要時間に基づいて第2の通信所要時間を検出することにより、比較的簡単に第2の通信所要時間が得られる。

【図面の簡単な説明】

【図1】実施の形態の負荷分散システムを備えるコンピュータネットワークの1ドメインの構成例を示すブロック図である。

【図2】実施の形態の負荷分散システムを備えるコンピュータネットワークの2ドメインの構成例を示すブロック図である。

【図3】自ノード負荷情報管理モジュールの負荷通知処理の内容を示すフローチャートである。

10 【図4】最適サーバ検索モジュールの検索処理の内容を示すフローチャートである。

【図5】ドメイン内の距離測定処理の内容を示すフローチャートである。

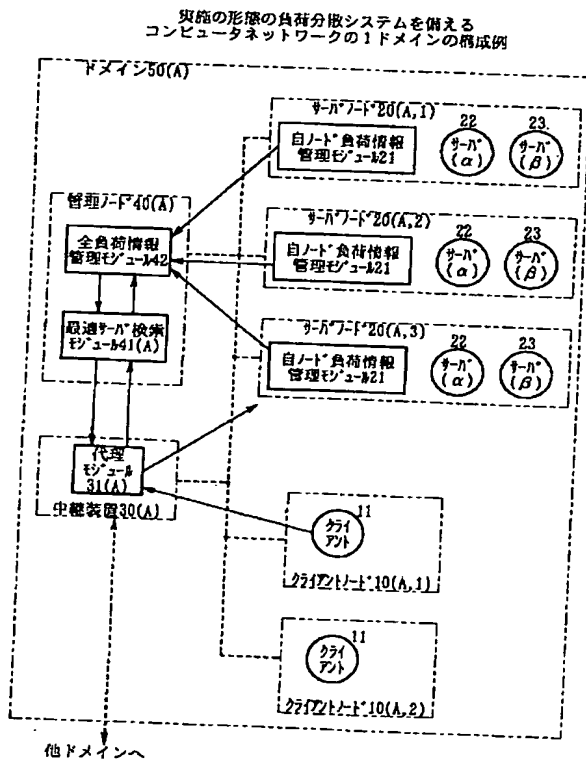
【図6】ドメイン間の距離測定処理の内容を示すフローチャートである。

【図7】従来の負荷分散システムの構成例を示すブロック図である。

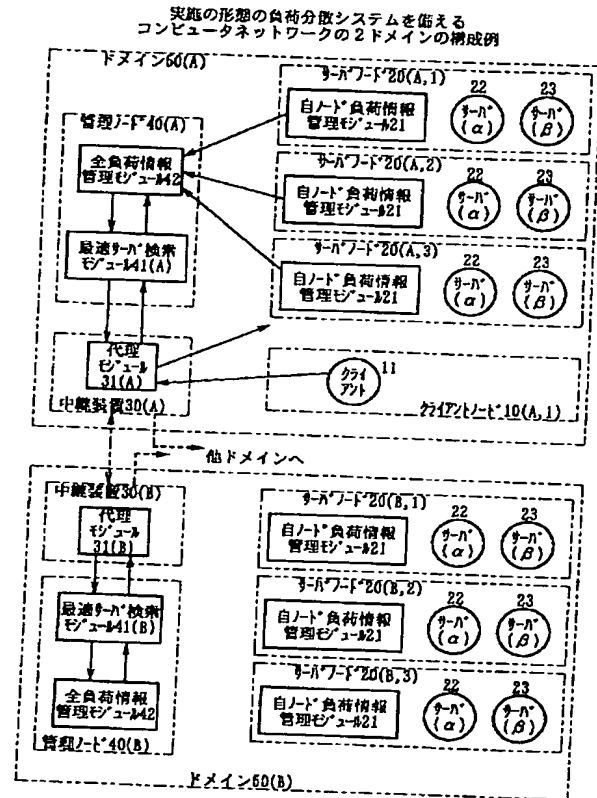
【符号の説明】

- 10 クライアントノード
- 11 クライアント
- 20 サーバノード
- 21 自ノード負荷情報管理モジュール
- 22, 23 サーバ
- 30 中継装置
- 31 代理モジュール
- 40 管理ノード
- 41 最適サーバ検索モジュール
- 42 全負荷情報管理モジュール
- 50 ドメイン

【図 1】

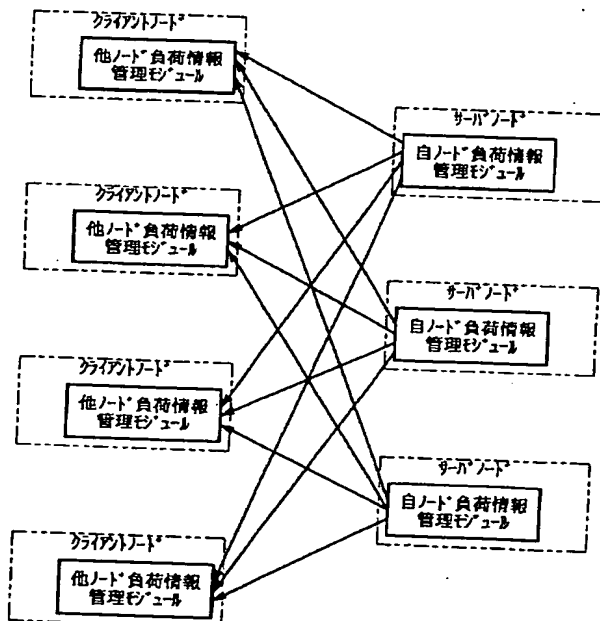


【図 2】



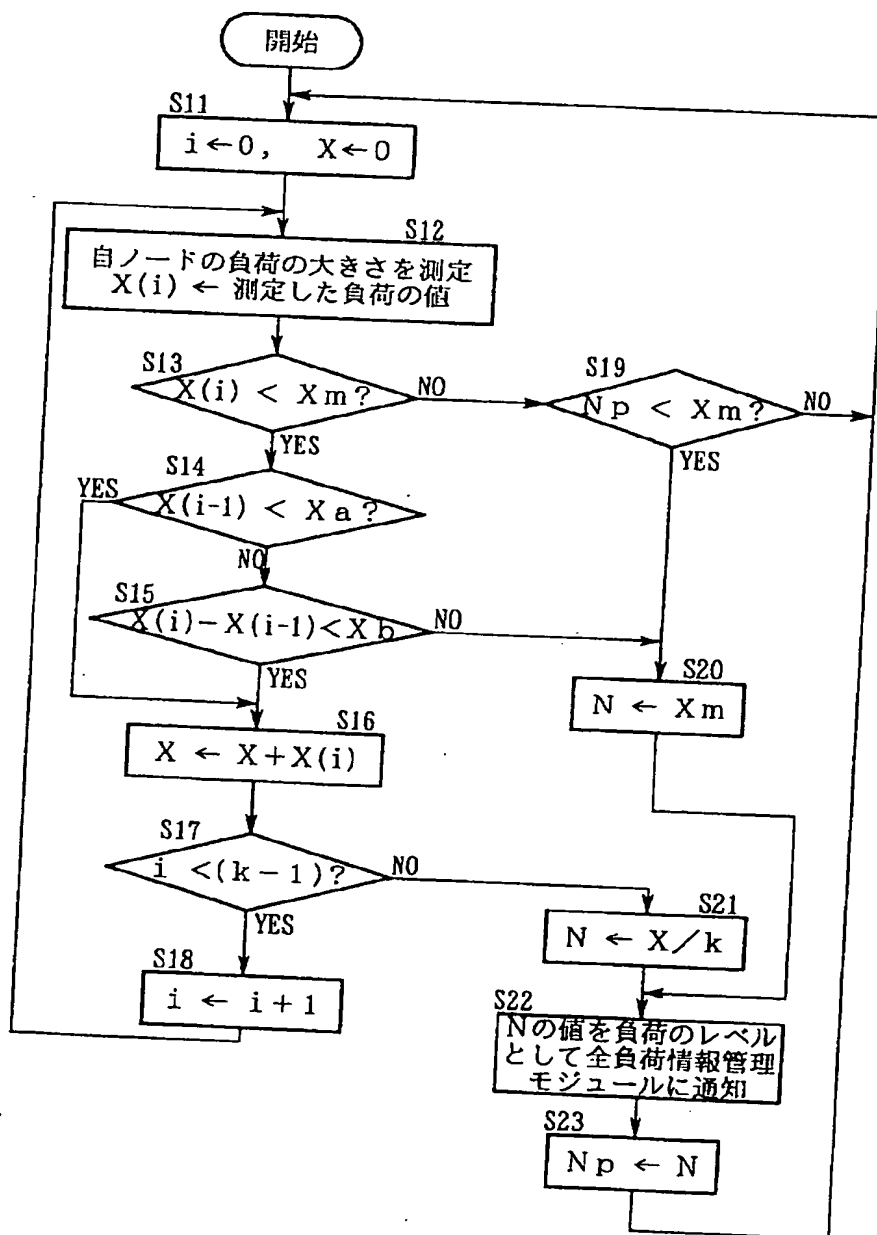
【図 7】

従来の負荷分散システムの構成例



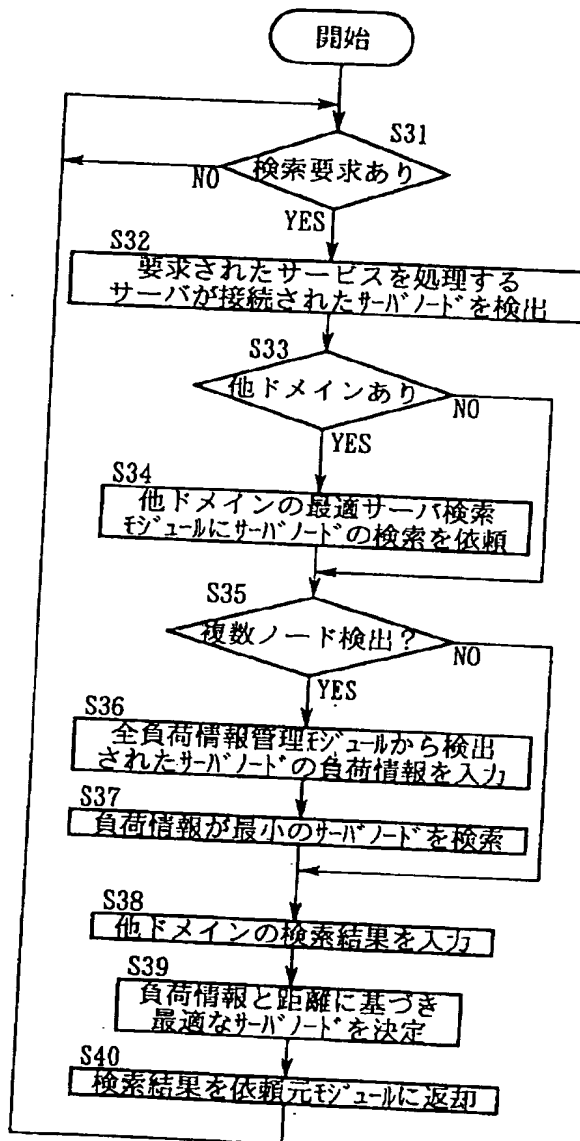
【図 3】

自ノード負荷情報管理モジュールの負荷通知処理の内容



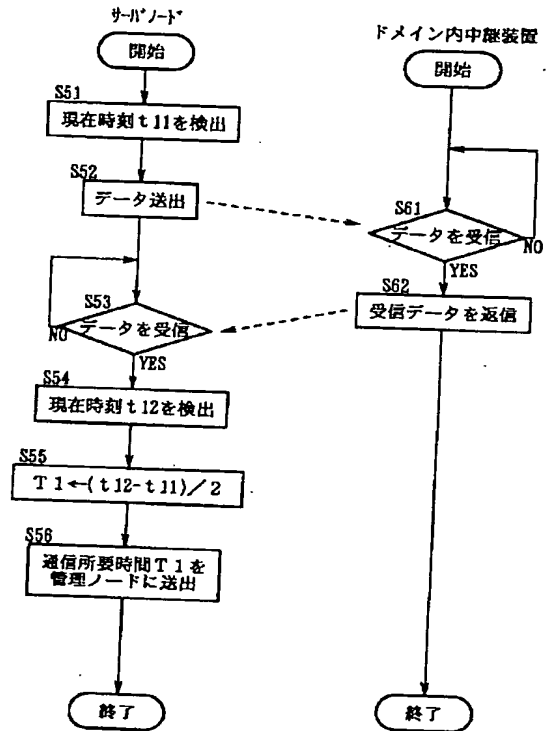
【図4】

最適サーバ検索モジュールの検索処理の内容



【図5】

ドメイン内の距離測定処理の内容



【図 6】

ドメイン間の距離測定処理の内容

